HOKKAIDO UNIVERSITY

Title	応用倫理学的探究において現状維持バイアスは排除されるべきノイズか
Author(s)	奥田, 太郎
Citation	応用倫理, 1, 15-29
Issue Date	2009-03
DOI	https://doi.org/10.14943/ouyourin.1.15
Doc URL	https://hdl.handle.net/2115/51661
Type	departmental bulletin paper
File Information	02_okuda_oyorinri_no1.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

応用倫理学的探究において現状維持バイアスは排除されるべきノイズか

奥田太郎（南山大学）

はじめに

応用倫理学が取り組む問題は多種多様であるが、その中心部には、新たな科学技術の社会的導入に際して重要な意味をもつ倫理的含意の剔出と検討という大きな課題が横たわっている。この課題は通常、新たな科学技術の導入によってわれわれの置かれている現状が何らかの影響を受けて変化する、という事態に向き合うこととして現れる。そして、多くの場合、批判の矛先は、現状から昔の状態に後戻りしようとする試みに対してではなく、現状を何らかの意味で「改善」しようとする試みに対して向けられることになる。そして、今まさにここで改善をカッコ付きで述べたように、総合的に勘案された結果それは実際には改善とは呼べないのではないか、という懸念がしばしば倫理的懸念として表明され、結果として現状維持が支持されるに至る。本稿で議論の俎上に置きたいのは、こうした類いの懸念の妥当性である。そうした懸念は、人間心理に深く根ざす現状維持へのバイアスによって形成されたものにすぎないのか否か。また、応用倫理学的探究に携わる論者の依拠する倫理的直観もまたその種のバイアスを抱え込みうるのだとすれば、そうしたバイアスは排除されるべきノイズだとみなされるべきか否か。本稿では、これらの問いに応えることを目指して、まずは合理的意思決定に関する心理学領域での成果、すなわち、カーネマンとトヴェルスキーのプロスペクト理論に基づく現状維持バイアスの研究についてごく簡単に述べ、その後、ボストロムとオードによる現状維持バイアス除去推進論、および、プロスペクト理論に対する倫理学の独自性を主張するフランシス・カムによる議論を考察の里程標とする。これらを経て、プロスペクト理論の明らかにした現状維持バイアスと、倫理学の取り扱う倫理的直観とは慎重に切り離して考えられるべきである、と結論し、表題の問いに対する現時点での回答を提示したい。

1. プロスペクト理論における現状維持バイアス

米国の心理学者ダニエル・カーネマンとエイモス・トヴェルスキーは、従来の経済学や合理的意思決定理論が前提としてきた、合理的人間像に依拠する期待効用理論の行動記述モデルとしての欠陥を経験的に明らかにした上で、それに代わる新たな行動記述モデルとしてプロスペクト理論（prospect theory）を提示した。彼らの理論は、行動経済学という領域の開拓を導き、2002 年にカーネマンはその功績によりノーベル経済学賞を受賞した[1]。このプロスペクト理論については、カーネマン自身が編集したアンソロジー（Kahneman & Tversky 2000）にて詳しく知ることができるので、理論の詳細はそちらを参照していただくとして、本稿では、現状維持バイアス（Status Quo Bias）に関連する部分に焦点を絞ってその概略を

[1] コロブキンによれば、法学関連の学術誌に掲載された論文のうち、「授かり効果」や「現状維持バイアス」のどちらかの語を用いていたものは、1990 年で 2 本、2001 年には 67 本、2003 年 1 月になると 373 本に増加している。（Korobkin 2003, p. 1227）

見るにとどめたい。

プロスペクト理論

多くの実験結果の積み重ねからカーネマンらが到達したプロスペクト理論の要点は、以下の如きである。(1) 人びとは一般に、利得よりも損失を大きく評価する。(2) 評価は一般に、相対的な準拠点からの変化によって形成される。(3) 人びとは一般に、確率の高い事柄を過小評価し、確率の低い事柄を過大評価する。(4) 人びとは一般に、利得についてはリスク回避を目指し、損失についてはリスク追求を目指す。これらは、リスク下での意思決定に関する一般的な心理的傾向を述べたものだが、このうちの (1) と (2) については、リスクが問題にならない取引状況においても同様の分析結果が得られる[2]。

こうしたプロスペクト理論のテーゼは、カーネマンらによる既存の期待効用理論の見直しに由来している。1980 年代前半に、カーネマンとトヴェルスキーは、合理的選択理論が依拠する効用に関する原理のうち、すでに多く批判の的とされていた推移性 (transitivity) と代替性 (substitution)[3]のみならず、多くの理論が取り入れている優越性 (dominance) と不変性 (invariance) の原理でさえも、維持するのが困難になるケースが多く存在することを経験的な研究によって示した (Kahneman & Tversky 2000, pp. 4-7)。優越性とは、見込み A が少なくともあらゆる点で見込み B と同程度によく、かつ、少なくとも一つの点で B よりもよいならば、A は B より選好されねばならない、という原理であり、不変性とは、複数の見込みの間の選好順序は、それらの見込みが記述される仕方に依存してはならない、という原理である。

たとえば、1981 年に *Science* 誌で公表 (Tversky & Kahneman 1981) されて以降、さまざまな論文の中で幾度となく言及されてきた有名な「アジア風邪の実験」では、選好の不変性が維持されない。概略は以下の通りである。

> 問題 1：米国で強力なアジア風邪が大流行し、600 人が死亡すると予想される。これに対抗するべく提案された方策は 2 つであり、それぞれの方策の結果は科学的に以下のように見積もられている。もし方策 A が採用されれば、200 人が助かる。もし方策 B が採用されれば、1/3 の確率で 600 人全員が助かり、2/3 の確率で誰も助からない。二つの方策のうちどちらを支持するか？
>
> 問題 2：問題 1 と同様の設定で、もし方策 C が採用されれば、400 人が死ぬ。もし方策 D が採用されれば、1/3 の確率で誰も死なずにすみ、2/3 の確率で 600 人全員が死ぬ。二つの方策のうちどちらを支持するか？

---

[2] 「多くの意思決定問題は、現状に留まることと、ある点では有利だが他の点では不利であるようなそれ以外の選択肢を受け入れることとの間での選択という形をとる」ので、現状 (status quo) を準拠点とみなし、選択肢の利点と欠点を利得と損失という形で捉えることで、取引状況における損失回避と現状維持の傾向を看て取ることができる。(Kahneman & Tversky 2000, p. 13)

[3] 推移性とは、A が B より選好され、B が C より選好されるならば、A は C より選好される、という原理であり、代替性とは、A が B より選好されるならば、A か C かのどちらかが五分五分で得られる見込みが、B か C かのどちらかが五分五分で得られる見込みより選好される、という原理である。

問題 1 では、回答者 152 名のうち、72%が A を選び、28%が B を選んだ。それに対して問題 2 では、回答者 155 名のうち、22%が C を選び 78%が D を選んだ。方策 A と方策 C、および、方策 B と方策 D は、フレーミングの仕方が異なっているだけで、それぞれの帰結は同じである。それにもかかわらず、問題 1 では方策 A が選ばれ、問題 2 では方策 D が選ばれる強い傾向があることがわかった。この傾向は、同一の回答者に数分以内で両方の問題に取り組ませても、問題文を読み直させてから回答させても、消えることがなかった。

この実験からわかるのは、問いのフレーミングが、助かる＝利得（gain）を軸になされているか、死ぬ＝損失（loss）を軸になされているかに応じて、問題に対する回答に正反対の傾向がみられる、ということである。すなわち、選好の不変性が維持されていないということである。

また、別のケースでは、選好の優越性が維持されないことがわかる。

問題 3：どちらかを選んで下さい。

E. 25%の確率で 240 ドルが得られ、75%の確率で 760 ドルが失われる。

F. 25%の確率で 250 ドルが得られ、75%の確率で 750 ドルが失われる。

この問題に対しては、86 名の回答者の 100%が F を選択した。選好は明らかに F が E に優越している。

問題 4：次の 2 組の決定をしなければならないとして、まず両方を検討した上で、自分の選好する選択肢を挙げて下さい。

決定 1

A. 240 ドルを確実に得られる。

B. 25%の確率で 1000 ドル得られるが、75%の確率で何も得られない。

決定 2

C. 750 ドルを確実に失う。

D. 75%の確率で 1000 ドル失うが、25%の確率で何も失わない。

この問題では、決定 1 に対して、150 名の回答者中の 84%が A を選び、16%が B を選んだ。決定 2 に対して、13%が C を選び、87%が D を選んだ。回答者のうち、A と D を選んだのは 74%、B と C を選んだのは 3%だった。

ところが、A に D を加えると、25%の確率で 240 ドルが得られ、75%の確率で 760 ドルが失われることになる。これは見込みとして問題 3 の E に相当する。他方、C に B を加えると、25%の確率で 250 ドルが得られ、75%の確率で 750 ドルが失われることになる。これは見込みとして問題 3 の F に相当する。

この実験からは、先ほどと同様に、問いのフレーミングが利得と損失、確実性と不確実性との組み合わせについて異なることで、見込み E と見込み F の間の選好の優越性が維持され

ない、ということがわかる。

しかしながら、期待効用理論の前提となる理想的な合理的人間像がわれわれの実像からかけ離れているというこの種の批判は、実際の現象をより見通しよく表現し理想的状況下での将来予測をするという理論の役割にとってそれほど致命的ではない。むしろ、カーネマンとトヴェルスキーの研究の刮目すべき点は、われわれの選択に潜む基本的な心理的傾向を実験結果の積み重ねによって示しただけでなく、その傾向を取り込んだ上で、効用理論をより実像に近い仕方で独自に形式化し、その予測精度を上げたところにあるだろう。カーネマンとトヴェルスキーは、いわば、上記のような「フレーミング効果」を意思決定におけるノイズとしてそれを排除する方向に向かうのではなく、それに正面から向き合って背後にある一般的な傾向をつかみとったのである。

さて、カーネマンらの業績の中で、本稿の目的に照らして重要だと思われるのは、「授かり効果（endowment effect）」および「現状維持バイアス（status quo bias）」、そしてそれらを説明しうる基本的な心理的傾向としての「損失回避（loss aversion）」である。以下、この三者についてコンパクトに論述したカーネマンらの論考（Kahneman, Knetsch & Thaler 1991）を参考に概説する。

授かり効果

「授かり効果」とは、リチャード・セイラーが 1980 年の論文（Thaler 1980）において、人びとが一度手にしたものを手放すのを厭うさまを記述するのに作り出した用語である。セイラーによれば、財産を手に入れる喜びよりもそれを手放す苦痛の方が大きいとき、買値は売値よりも顕著に低くなる。つまり、ある人が財産を手に入れるために支払う最高価格は、その人が一度手にした財産を手放す代償として求める額の最小値よりも少なくなる。代表的な実験としては、被験者にマグカップとお菓子を手渡し、それらを交換する取引を行わせる、というものである。マグカップとお菓子の価格にかかわらず、被験者はそれぞれ自分が最初にわたされた方のアイテムを保持する傾向にある、というのがその結果である[4]。

---

[4] 授かり効果を検証するためのマグカップ実験は、きれいな空気や環境へのダメージなどを測る仮想評価法（Contingent Valuation Method）では WTP と WTA にギャップが生じる、ということからインスパイアされてもいる。（Kahneman & Tversky 2000, p. 294-295.）

この効果について、カーネマンらは、市場の中で経験を積めば減少する類いのものか否かを確かめるための複数の実験を行っている（Kahneman, Knetsch & Thaler 1990）。たとえば、コーネル大学の学部生向けの経済学教室で学生に市場取引に関する実習をさせた後で、マグカップとボールペンの交換に関する実験に参加させた。実験は、被験者を、マグカップを与えられた売り手役とそれを購入する買い手役に分けて、それらの取引を行わせるものであり、複数回の取引の中で、前の取引での状況に関する情報が共有されるような仕組みを導入して、継続的な取引の中に学習の要素を組み込むように工夫が施されたものである。結果として、売値の平均値は買値の平均値の約二倍であり、取引量は標準的な経済学理論の予測の半分以下であった。要するに、学習と経験の機会があったにもかかわらず、多くの学生が一度手にしたものを手放すことを避けた、ということである。

さらに、上記のように取引量が少ないのは、売ることへの躊躇に由来するのか、買うことへの躊躇に由来するのか、を確かめる実験が、サイモン・フレーザー大学の 77 名の学生を対象に実施された。被験者を三つのグループに分け、売り手グループにはマグカップを与え、0.25 ドルから 9.25 ドルまでの値段で売るように求め、買い手グループには同様の値幅で買うように求めた。選び手と名付けられた第三のグループにはマグカップを与えず、同様の値幅を提示して、マグカップを受け取るか対応する額のお金を受け取るかを選ぶように求めた。売り手と選び手は客観的には同じ状況（設定した値幅でマグカップを選ぶかお金を選ぶか、という状況）に置かれている。にもかかわらず、選び手はむしろ買い手によく似た振る舞いをした。留保価格の平均値は、売り手 7.12 ドル、選び手 3.12 ドル、買い手 2.87 ドルであった。ここから示唆されるのは、

現状維持バイアス

「現状維持バイアス」は、サミュエルソンとゼックハウザーが 1988 年の論文（Samuelson & Zeckhauser 1988）で、人びとが準拠点としての現状（status quo）を維持する傾向があることを記述するために用いた用語である。彼らが行った実験は、基本的には以下のようなパターンのものである。

> 問題 5：あなたは新聞の経済欄の熱心な読者だが、投資の初心者である。今度大叔父から多額の財産が相続されることになった。あなたの選択できる投資先は、ほどほどのリスクの企業、リスクの高い企業、国債、地方債である。あなたはどれを選ぶか？
>
> 問題 6：あなたは新聞の経済欄の熱心な読者だが、投資の初心者である。今度大叔父から現金と証券が相続されることになった。そのうちの大部分はほどほどのリスクの企業にすでに投資されている。あなたの選択できる投資先は、ほどほどのリスクの企業、リスクの高い企業、国債、地方債である。あなたはどれを選ぶか？
>
> 問題 5 は、「現状」を定めないニュートラルなケースであり、問題 6 は、「現状」をあらかじめ定めてある。

こうしたデザインの実験をさまざまなシナリオで複数試みた結果、共通していたのは、選択肢は、それが「現状」である場合、あるいは、ニュートラルなケースで選ばれる頻度が高く「現状」の代替選択肢として遜色のない場合、選ばれる確率が高いということであった。いずれにせよ、一連の実験からわかったのは、選択肢は「現状」として設定されると顕著に好まれるようになった、ということ、および、「現状」の優位は選択肢の数とともに増大するということ、であった。

このバイアスについては、米国で起こった実例がある。1980 年代に、ニュージャージー州とペンシルバニア州で、自動車保険について、廉価だが訴訟の権利が限定されているものと、高価だが無制限に訴訟の権利が認められるものとを選ばせることになった。ニュージャージー州では、高価な保険がデフォルトとして提示され、ペンシルバニア州では、廉価な保険がデフォルトとして提示された。その結果、両州で、多くのドライバーが、デフォルトとして提示された方の保険を選ぶことになった。（Kahneman & Tversky 2000, p. 240, 294） これは、政策決定に及ぼす現状維持バイアスの影響を示す古典的実例とされている。

損失回避

授かり効果、現状維持バイアスの研究から得られる一般的な結論は、（1）効用の重要な担い手は、富や福利の状態ではなく、準拠点に相対的な変化であり、（2）事態を悪化させる変化（損失）の方が、事態の改善（利得）よりも大きく見積もられる、ということである。カ

取引量が低いのは、買い手のお金を手放すことへの躊躇ではなく、売り手（マグカップの持ち主）の持ち物（endowment）を手放すことへの躊躇に由来する、ということである。

ーネマンらは、授かり効果と現状維持バイアスは、より基本的な概念である「損失回避」によってうまく説明できる、と考える（Kahneman & Tversky 2000, pp. 165-168.）。現状維持との関係で言えば、現状を変更して得られる利得を逃すことよりも、現状を変更して生じる損失をより強く忌避するので、結果として現状維持が選択される。一般に、損失回避によって変化よりも安定が好まれる、というわけである[5]。

さらに、カーネマンらが 1986 年に発表した論文によると（Kahneman, Knetsch & Thaler 1986）、トロントとバンクーバーの住人を対象に電話調査を実施したところ、フェアネスに関する判断に損失回避の傾向が強く見られた。その調査において一連の質問の中に埋め込まれた関連質問は、たとえば以下のようなものであった。

> 問題 7a：ある会社が小さな利益を上げた。その会社のある地域では、不景気で実質的な失業が生じているがインフレは生じていない。その会社は、今年の賃金を 7%下げることにした。
>
> 回答者数 125　受け入れ可能（37%）　アンフェア（63%）

> 問題 7b：ある会社が小さな利益を上げた。その会社のある地域では、不景気で実質的な失業が生じており 12%のインフレが生じている。その会社は、今年の賃金を 5%上げることにした。
>
> 回答者数 129　受け入れ可能（78%）　アンフェア（22%）

問題 7a と 7b はいずれも、実質的には賃金が 7%下がることになり、富や財産に関する帰結としては同じなのだが、人びとは一般に、それを損失と捉える（現状が準拠点となり 7%の賃金引き下げは損失として捉えられる）か、利得と捉える（起こってしまった 12%のインフレが準拠点となり 5%の賃金引き上げは利得として捉えられる）かで、アンフェアか否かの判断を変化させる傾向にあるようである。フェアネスに関する判断は、一種の倫理判断であり、当然ながら、倫理学者はこうした言説の流通に対して無頓着でいるわけにはいかないだろう。この問題については、本稿 3 にて論じることにする。

以上、カーネマンらのプロスペクト理論、とりわけ、損失回避に牽引される授かり効果と現状維持バイアスについて概観してきた。次節からは、これらに対する倫理学者からの応答を検討する。

2. 反転テスト

まずは、現状維持バイアス除去推進論をとるボストロムとオードの試み（Bostrom & Ord 2006）を見てみよう。彼らによれば、われわれの意思決定が基づく判断は、数々のバイアス

---

[5] この損失回避の傾向は、同じ額を手放すにもかかわらず、それが損失とみなされるかコストとみなされるかによってわれわれの判断が変わってくる、という仕方でも現れる。たとえば、賭けへの参加料として 5 ドル支払って、その賭けにおいて 10%の確率で 100 ドルが得られ 90%の確率で何も得られない、という場合の方が、賭けにおいて 10%の確率で 95 ドルが得られ 90%の確率で 5 ドルを失う、という場合よりも容易に受け入れられる傾向がある（Kahneman & Tversky 2000, p. 15）。

にさらされており、そうしたバイアスを自覚したり取り除いたりすることは、時として、われわれの判断を大きく改善する。彼らは、その種のバイアスの自覚と除去には、心理学や社会科学からの経験的な情報が有益でありうるとして、応用倫理学的問題における現状維持バイアスを除去するテストを提案するに至る。

ボストロムとオードは、たとえば、マグカップ実験が示しているのはわれわれの心の動きの事実であり、それ自体は不適切でも不合理でもない、と述べ、また、全面的に心理学研究の成果を支持するわけではないが、ひとまず現状維持バイアスが広くわれわれの心理に見られるということを前提として議論を進める、という慎重な姿勢を論文冒頭で強調している。その上で、彼らはまず、「現状維持バイアス」を（カーネマンらのそれとは異なる意味をもたせて）「現状を不適切に好むこと」と定義する（Bostrom & Ord 2006, p. 659）。そして、現在、しきりに議論されている認知能力のエンハンスメントを検討事例として掲げ、主として、帰結主義的議論の土俵で論述を進めて行く。ボストロムとオードによれば、認知能力のエンハンスメントをめぐる論争において、賛否両陣営は互いの直観のバイアスを難詰しているように思われる。それゆえ、どちらの直観に問題があるのかを裁定する手段が必要であると考えられ、彼らはその発見法として「反転テスト（Reversal Test）」を提案する。

> **反転テスト**：あるパラメータを変更しようという提案が全般的に見て悪い帰結をもたらすと考えられる場合には、同じパラメータを逆方向に向けて変更することを考慮せよ。もしその変更もまた全般的に見て悪い帰結をもたらすと考えられるのであれば、なぜそのパラメータの変更によってわれわれの状況が改善されえないのかについて説明する責任は、それらの結論に至った者の側にある。もし彼らがそのことを説明できなければ、彼らが現状維持バイアスの下にあると疑うだけの理由があることになる。（Bostrom & Ord 2006, pp. 664-665）

これは、現状維持を主張する者には、現状が局所的に最適である理由を説明する責任があると考え、その理由の提示の可否を見るテストである。たとえば、認知能力を増強すると現状より悪い帰結が生じる、と主張する者が、同時に、認知能力を減退させても現状より悪い帰結が生じる、と主張するならば、彼らは現状の認知能力が局所的最適であるとみなしていることになる。そのことの証明は彼ら自身が行わねばならない。それができないなら、彼らは現状維持バイアスの下にあると診断されるのである。

ボストロムとオードは、さらにこのテストを補強するべく、次のような仮想事例の検討を要求する。

> 仮想事例：唯一の水源が汚染され、汚染の影響でわれわれの認知能力が低下することがわかった。幸いなことに、安全な体細胞遺伝子治療によってわれわれの認知能力を増強して、汚染で低下した認知能力を「現行水準」まで高めることができるので、すべての人にその治療が施されることになった。数年後、汚染物質がなくなり、それが脳に与えていたダメージが消えるにつれて、われわれの認知能力の低下はなくなってきたが、そ

> うすると、遺伝子治療による認知能力の向上のせいでわれわれの認知能力は「現行水準」よりも高くなってしまうことがわかった。この場合、再び遺伝子治療によって認知能力を「現行水準」まで落とすべきか、それとも汚染物質を意図的に摂取し続けて認知能力の「現行水準」を保つべきか。もし向上した認知能力を保つのであれば、それは事実上、「現行水準」からのエンハンスメントになる。この選択は奇妙には思われない。では、汚染が発生しない場合の認知能力に対するエンハンスメントとどう異なるのか。（p. 672）

この仮想事例に含まれる議論構造を一般化すると以下のような発見法が得られる。

> **二重反転テスト（Double Reversal Test）**：あるパラメータを増やすことと減らすこととの両方が全般的に見て悪い結果をもたらすと考えられる、と想定する。ある自然の要因がパラメータを一定の方向に動かす恐れがあるというシナリオを考え、現状を維持する介入によってこの変化の埋め合わせをすることが善いかどうかを問うてみる。もし善いのであれば、その後に、かの自然に生じた要因が徐々に消滅していくと考えて、最初の介入を反転させる介入を行うことが善いかどうかを問うてみる。もし善くないのであれば、自然な相殺要因がない場合であっても最初の介入を行うのは善いことであると考える一見自明な強い言い分がある。（p. 673）

このテストによって俎上に載せられるのは、二つの「現状」である。一つは、（a）当該パラメータの現在の平均値であり、もう一つは、（b）介入しない場合にもたらされるデフォルト状態の値である。（a）の現状を守るなら、汚染に由来する能力低下を補正するエンハンスメントをすることになり、また、（b）の現状を守るなら、汚染の影響が消えた際に逆向きエンハンスメントを差し控えることになる。ボストロムとオードによれば、これらの二つの「現状」を対照させることで、現時点のパラメータを変更した場合に予想される利益についてわれわれがもつ直観に現状維持バイアスが影響しているかどうかを突き止めることができる。（ibid.）

なお、ボストロムとオードは、この二重反転テストを用いて、身長のエンハンスメントには現状維持バイアスに左右されない差し控えるべき理由を見出しうる、と考える。身長は現状から高くなっても低くなっても、建物や衣類等のすでにあるインフラを変更するコストが大きいので、現状の身長が局所的最適であり、身長のエンハンスメントはなすべきではない、ということになる[6]。また、彼らは、マイケル・サンデルによる「招かれざるものに開かれてある（open to the unbidden）」ことはわれわれにとって善いことである、という理由に基づく遺伝子エンハンスメント反対論（Sandel 2004）[7]を採り上げてもいる。この場合は、親の

[6] ただし、身長のエンハンスメントと言うとき想起されるのは、巨人や小人になることではなく、現時点で「背が高い」と言われている人程度の背の高さへのエンハンスメントであろうから、ボストロムとオードのここでの議論が身長エンハンスメント全般に適用可能であるとは言い難い。

[7] サンデルの *The Atlantic Monthly* 誌上に掲載されたこの小論は、後に加筆・修正され、著書（Sandel 2007）として公刊されている。

子孫に対する影響力が現状よりも小さくなればなるほど（招かれざるものにより開かれてあるがゆえに）善い、ということになり、現状の局所的最適を主張してはいないため、反転テストをパスすることができるだろう、と彼らは述べている。

考察

ボストロムとオードは、認知能力のエンハンスメントなど新たな科学技術の導入に際する応用倫理学的探究において、論争をいたずらに長引かせる要因となる「現状維持バイアス」は排除されるべきノイズであると位置づけ、現状が局所的に最適か否かを判定するクリティカルシンキングが必要である、という合理主義的な立場で議論を進めている。そして彼ら自身が明言しているように、反転テストは、あくまでも帰結主義的考慮のためのツールであって、義務論的議論の中で用いる際には、その帰結主義的側面に対してのみ適用されうるものであり、彼らの議論の射程は、帰結主義の内部に留まる。したがって、反転テストは、応用倫理学的課題に対して合理主義的に臨むのが望ましいとする帰結主義者が自身の「現状維持バイアス」を排除するためには有効に機能するだろうし、また、現在の認知能力レベルの局所的最適性が論点となる場面では効力を発揮するだろう。こうした方向でのわれわれの判断の洗練は、現在直面している事態をより正確に捉えるために欠かすことのできない作業であることはまちがいない。

しかしながら、問題は、応用倫理学的探究から「現状維持バイアス」を排除すると謳う彼らの議論が、結局は、帰結主義的判断を洗練するもの以上のものではなく、われわれの倫理的直観と現状維持バイアスの関係という核心的な問いにまったく答えうるものではない、というところにある。

たとえば、ボストロムとオードは、仮想事例における二つのシナリオ（汚染あり／なし）には重要な違い（たとえば、為すこと（doing）／許すこと（allowing）の違い）がある、とする非帰結主義的な反論を想定して、それはエンハンスメントの是非という問題の根本には無関係な反論である、と応酬している。というのも、シナリオは違えども、エンハンスメントが施された状態が実現されていることに変わりはないからである。確かに、自然の要因に対して抗うために施された処置が、結果としてエンハンスメントになってしまうことは道徳的に認められるが、ただ能力を高めることだけを意図してエンハンスメントを施すことは認められない、と述べる非帰結主義的な論法では、エンハンスメントの根本的禁止ではなく文脈的・一時的禁止を主張するに留まるであろう。しかしながら、エンハンスメントをめぐる帰結主義者と非帰結主義者の倫理学的議論のポイントは、現状の認知能力レベルの局所的最適性にあるというよりも、科学技術の導入のプロセスのありようにこそある、ということもできよう。そうだとすれば、ボストロムとオードの提案は、一定の示唆を与えてくれるとはいえ、あまりに限定され過ぎたものであると言わざるを得ない。

また、彼らの現状維持バイアスの定義「現状を不適切に好むこと」は、言い換えれば、現状を局所的最適点と誤ってみなすこと、ということになるが、これは、カーネマンらの現状維持バイアスに関する研究成果とは何の関係もない。カーネマンらの言う「現状」とは、われわれが判断（あるいは効用形成）に際して依拠する準拠点であり、そこに固執するのは、

損失回避という心理的傾向の結果なのであって、局所的最適か否かの判断が誤っているからではない。カーネマンら自身が、フレーミング効果は判断の誤りではなく錯視などの知覚の誤りに類するものである、と述べているように（Kahneman & Tversky 2000, p. 5）、バイアスそれ自体の消去は不可能に近いとみなした上で、判断の誤りについて吟味する、という慎重な態度がより望ましいであろう。ボストロムとオードの提案は、カーネマンらの業績とは独立に成立しうるものなのだから、いたずらに心理学と倫理学を接続させるかの如き印象を与えるのは慎まれるべきである。

他方で、仮にカーネマンらの路線を推し進めて倫理学に切り込むのであれば、主として義務論者が依拠する為すこと／許すことの区別等に関するわれわれの倫理的直観が、すべてプロスペクト理論の枠組みの中で説明される、とする筋がありうるだろう。次の節では、そうした筋の議論に対抗するべく提示されたカムの議論を検討する。

3. 倫理学における現状維持バイアスの位置づけ

フランシス・カムは、近年、自らの義務論的倫理学理論のまとまった仕事（Kamm 2007）を公刊したが、その一章として、プロスペクト理論と義務論的倫理学との関係を細かく取り扱った論文（Kamm 1998）が収録されている。カムの批判の矛先は、現状維持バイアスなどの心理的傾向に対してではなく、カーネマンらがそうした傾向を読み取ってきた元の実験データとその解釈それ自体に対して向けられている。

カムにとって何よりも重要な問いは、カーネマンらの実験データによって、害をなすこと（harming）／手助けしないこと（not-aiding）の区別の重要性が失われるか、であった。為すこと／許すこと、害をなすこと／手助けしないこと等の区別は通常、それ以上別の根拠に遡れないわれわれの倫理的直観によってなされるものだと考えられ、それゆえ、義務論者の強力な理論装置として用いられてきた。もしこうした直観による区別が、実は損失回避という心理的傾向に牽引されたバイアスに基づくものにすぎないということになれば、その倫理的な効力のみならず基本原理としての座をも失ってしまうだろう。これは、とりわけカムら義務論者にとっては避けて通れない由々しき事態だったのである。

そこでカムは、プロスペクト理論における「損失／無利得（loss/no-gain）」の区別は、非帰結主義的理論における「害をなすこと／手助けしないこと（harming/not-aiding）」の区別と同一か、と問う（Kamm 2007, p. 430）。というのも、害をなすことは誰かに損失を与えることであり、手助けしないことは誰にも利得を与えないことである、と定めれば、プロスペクト理論の枠内で非帰結主義的な区別が処理できるように思われるからである。

これについてカムは、まず以下の 3 点を指摘する。（1）負わされた損失（imposed loss）と損失それ自体（loss per se）とは異なり、かつ、負わされた無利得（imposed no-gain）と無利得それ自体（no-gain per se）とは異なる。（2）害をなすこと／手助けしないことの区別は、被害者が受けるもののみならず行為者のなすことやなさぬことに関わるが、損失／無利得の区別は、被害者が受けるものにのみ関わる。（3）プロスペクト理論による損失の記述の仕方にしたがえば、手助けをしないことで損失が生じることがありうる。

このうち特に（3）について議論が進められる。プロスペクト理論では、損失を被ること

（suffering losses）と利得を受けられないこと（not receiving gains）との間にのみ決定的な区別を設けるので、損失を被るのを見過ごすことと、損失を引き起こすこととの区別をすることができない。そうした区別を等閑視するのであれば、たとえば、健康な人の命が失われるのを妨げないことに決めた場合であろうと、彼の命が失われるのを引き起こした場合であろうと、その人の命が失われるという損失が生じた同一の事態ということになる。これに対して、害をなすこと／手助けしないことの区別を重んじる倫理学理論では、それらの間には道徳的性質の上で重大な差がある、と考えるのである。

カムはさらに分析を続ける。害をなすこと／手助けしないことの区別は、ある人が、われわれの関与がなければもち続けたであろうものを損失すること（害をなす場合）と、われわれの関与があればもち続けられたであろうものを**損失する**こと（手助けしない場合）との間の区別として立ち現れる。さらに、われわれの関与がなければもち続けられたであろうものを損失すること（害をなす場合）と、われわれの関与があれば得られたであろうものを**得られなかった**こと（手助けしない場合）との間の区別としても立ち現れるのである。ここからは、害をなすこと／手助けしないことの区別と、損失／無利得の区別とがそれぞれ別の区分として成立しているのがわかる。手助けしないことを表現するフレーミングは損失から無利得へと切り替わっていても、それらの許容可能性に関する道徳判断は変化しない。すなわち、損失／無利得の場合に生じるフレーミング効果は、道徳判断には影響を及ぼさない、ということである（Kamm 2007, p. 431）

カムはさらに、フィリッパ・フットに由来する思考実験によって、これら二つの区別の隔たりを確認する（Kamm 2007, pp. 433-434）。

> 救助事例 1：われわれは、瀕死の 5 名を乗せて病院へと車を走らせている。路上に、死にかけの者が一人倒れていて、われわれ以外に彼女を救うことはできない。もし彼女を助けるために停車すれば、5 名は助からないであろう。この場合、彼女を見殺しにする（let the one die）ことは許される。

> 救助事例 2：われわれは、瀕死の 5 名を乗せて病院へと車を走らせている。道の真ん中で、元気な人が身動きのとれない状態にある。もし 5 名を助けるために病院へと進めば、予見される副次的結果（a foreseen side effect）として、彼女をひき殺してしまうだろう。この場合、われわれはそれをすべきではない。

カムは、2 つの救助事例にみられる判断の相違を損失／無利得の区別に基づいて説明可能だと論じたホロヴィッツの議論（Horowitz 1998）[8]を批判する。ホロヴィッツによれば、たとえば、救助事例 1 では、全員が瀕死であり、彼らの利得が問題になっているので、1 対 5

[8] ホロヴィッツは、ウォレン・クインによって提示された、為すこと／許すことの区別を支持するために直観に訴える事例が、クイン自身の現状維持バイアスによってもっともらしく見えている（つまり、為すこと／許すことの区別をわれわれは道徳的に直観している、というわけではない）ということを、プロスペクト理論に基づき論じている。また、これにファン＝ローヤン（van Roojen 1999）が反照的均衡の観点から応酬しているが、本稿ではこの軸の論争には立ち入らない。

の利得間の葛藤状況とみなしうる。この場合には、より多い利得が選好されるため、5 名を救うことになる。それに対して、救助事例 2 では、路上の 1 名はわれわれが車を走らせることで命を失うが、車の 5 名はもともと瀕死なので、たとえ助からなくても損失を被るわけではなく利得がないだけである。つまり、救助事例 2 は損失と無利得の葛藤状況とみなしうる。この場合には、損失回避が選好され路上の 1 名を救うことになる。通常上記 2 つの救助事例は、為すこと／許すことの区別によって説明されるのだが、このように分析すれば、損失／無利得の区別によってプロスペクト理論的に説明可能だということになるわけである。

これに対してカムは、上述の、損失／無利得と害をなすこと／手助けしないことの軸のずれを示すべく、事例を少し書き換えることで反論している。たとえば、救助事例 1 を、路上に元気な人がいるが、われわれと一緒にいなければその人は野生の動物に殺されてしまうだろう、というように書き換えると、路上の人は損失を被ることになる。つまり、この救助事例 1*は、損失／無利得の枠組みのもとにある。しかし、この場合には、プロスペクト理論に反して、損失の回避は無利得の回避を上回らず、害をなすこと／手助けしないことの区別に基づき、彼女を見殺しにすることは許される、と判断される。

また、救助事例 2 を次のように書き換える。車の 5 名は今のところ元気だが、もう一度予防薬を打たないと数時間後に街に蔓延する病によって死んでしまう。しかし、薬は病院でしか打つことができない。この場合、5 名は損失を被ることになり、救助事例 2*は、1 対 5 の損失同士の葛藤状況となる。ここでは、損失／無利得の区別がなく、損失の規模の大小が明白であるにもかかわらず、（損失を予防することによって）誰かを手助けするために、（予見される損失を引き起こすことによって）誰かに害をなすことは許されない、という考慮が働き、路上の人をひき殺すことは許されない、と判断されるのである。

これら 4 つの救助事例の検討を通じて、損失／無利得の区別と、害をなすこと／手助けしないことの区別が、基本的には独立の枠組みであることが明らかになったはずである。こうして、損失／無利得の区別によって、害をなすこと／手助けしないことの区別を説明しようとする試みを粉砕した後、カムはさらに踏み込んで、カーネマンらの実験データそれ自体が、人びとの直観がむしろ害をなすこと／手助けしないことの区別に従っていることを示している、という挑発的な議論を展開している。

たとえば、本稿 1 で紹介した問題 7a と 7b では、7a をアンフェアだと考え、7b をそうではないと考える傾向がある、ということであった。カーネマンらはこれを損失回避の現象として読み取ったわけだが、カムは、害をなすこと／手助けしないことの区別に基づく人びとの倫理的直観の表れを見る（Kamm 2007, p. 435）。すなわち、会社は 7a では損失を引き起こしているが、7b では、インフレという外的要因によって生じる損失を防ぐことができないだけである。人びとのフェアネスに関する判断は、このような分析に基づく解釈にしたがえば、倫理的直観を適切に反映しているとみなしうるのである[9]。さらに、カムは、カーネマンら

[9] 問題 7a、7b のフェアネスに関する判断が、回答者のアイデンティティ（たとえば、経営者、正社員、非正社員、地元住民、政策決定者など）によって異なる可能性を指摘することで、カーネマンらの解釈に留保を突き付けることもできるかもしれない。もちろん、カーネマンらの調査は、フェアネスに関わる問題を巧妙に他のダミー問題と組み合わせて、回答者の第三者的立場からの判断を引き出すようにデザインされ、同一回答者に対比問題を出さない（たとえば、7a と 7b の一方のみを回答者に出題する）ようにした上で統計的に処理されたものだと思われるので、標準的な心理学研究としてはさしあたりアイデンティティ依存の要

によるその他の実験の中にも、倫理的直観の表れを読み取り、記述の道徳的等価性の判定を倫理学理論の仕事のひとつと位置づけ、プロスペクト理論の成果と倫理学の領分とを丁寧に切り分けている（Kamm 2007, pp. 438-445）。

考察

カーネマンら自身がプロスペクト理論を記述モデルと位置づけている（Kahneman & Tversky 2000, p. 65）ように、プロスペクト理論の役割は、現実にわれわれがとる行動の記述および予測である。それゆえ、よりよい記述たりえているか、よりよい予測を提供できるか、といった規範性が含まれることはあっても、われわれがどう行動すべきかを示す行為規範の提供が目指されているわけではない。彼らは、損失を回避すべきだとも利得をもっと重視すべきだとも主張しないし、現状維持バイアスを除去すべきか否かといった判断も下さない。

カムが上記の議論において試みたのは、そうした記述理論と規範理論の間のねじれ具合を解きほぐすことであったと言えるだろう。カムは、プロスペクト理論が明らかにしてきた心理学的な成果と、われわれの倫理的直観とが、さしあたり無関係なものとして位置づけられることを義務論的アプローチで示したのである。

むすび

ここまでの議論より、本稿冒頭の問いにはどのように応えられるだろうか。現状維持を支持する倫理的懸念は妥当なものなのか、現状維持バイアスに基づく一種の錯視にすぎないのか。この問いに対しては、前節でのカムの議論を敷衍して一部応えることができるかもしれない。フェアネスと損失回避に関するカーネマンらの調査が示した人びとの回答傾向が十分に妥当な倫理的応答であったように、現状維持を支持しているからといってバイアスを強く懸念すべき理由があるわけではない。しかし当然ながら、現代の倫理的諸問題では、問題 7a や 7b に比べれば遥かに複雑で多様な事柄を考慮しなければならず、そこでわれわれの心理的傾向が判断の中に入り込んでくることはありうる。

たとえば、臓器提供に関する方式としてオプト・インとオプト・アウトがあるが、これらは、デフォルトとして提示された選択肢を選びやすいという現状維持バイアスと密接に関係するため、臓器不足を解消したいと考える者は、臓器提供をデフォルトとするオプト・アウトを推奨することが多い。しかしながら、すでにオプト・イン方式が導入されている社会で新たにオプト・アウト方式への転換を行った場合、現状維持バイアスの作用により、臓器不足が解消されないかもしれない。そもそも現状維持バイアスとは、評価形成は準拠点からの変化に依存する、という現象であるので、すでにオプト・イン方式のもとで臓器移植に関す

素をそれほど重く扱う必要はないだろう。とはいえ、ここには、回答者のアイデンティティの影響を現行の統計的処理によって非有意なレベルにまで引き下げることができるのか、あるいは逆に、回答者のアイデンティティの影響という問題系を考える上で統計的手法はどの程度有効か、といった方法論上の問いが横たわっており、これは統計学の哲学や心理学の哲学が取り組むべき課題であろうし、そこから応用倫理学的探究が得るところも少なくないだろう。また、プロスペクト理論の文脈からは大きく離れて、フェアネスの判断のような倫理的判断のありようが、どの程度判断者の政治的・社会的アイデンティティに左右されるのか、といった倫理学と社会学にまたがる課題も想起される。しかしながら、これらの問いは本稿のテーマからは少し外れるため、取り扱わない。なお、本註で述べた点は、『応用倫理』査読者からの示唆によって改めて気付かされたところであり、有益なコメントに感謝の意を表し、今後さらに考察を深めていきたい。

る準拠点が定まっていると、そこからの変化を嫌ってアウトの意思表明が増える可能性は否めない。そこに医療業界への不信といった別要因が加われば、損失回避の傾向が強まることも考えられるだろう[10]。

とはいえ、こうしたことによって、臓器移植や臓器不足に対する倫理的懸念それ自体が左右されるわけではない。また、「現状維持バイアス」に対する意味づけが、プロスペクト理論の論者とボストロムらとの間で大きく異なっていたことを忘れてはならない。応用倫理学的探究においては、「現状維持バイアス」という名で通用している事柄が、どのような文脈下で有効なのかを正しく把握した上で、議論全体のどのあたりで、バイアスとして事柄を把握すればよりよい見通しがえられるのかを探り、議論の風通しをよくすることこそが肝要である。したがって、表題の問いに対する現時点での回答は、否である。

なお、応用倫理学的探究に携わる論者自身もまた現状維持バイアスに基づいた主張をすることから逃れるのは不可能であろう。だからこそ、さまざまな準拠点をもつ、関心を異にする多くの者たちによるコラボレーションが必要となるのである[11]。ただし、ここに及んでも、応用倫理学的探究に携わる論者にとって現状維持バイアスは常にノイズとして除去されるべきである、ということにはならない。重要なのは、そこにバイアスが介在しうることを自覚し、それが問題の文脈の中でどのような意味をもっているのかを明確にすることであって、単にバイアスを消去することではない。バイアスがわれわれのよき生を妨げるとは限らないのである。

文献

Bostrom, Nick and Ord, Toby, “The Reversal Test: Eliminating Status Quo Bias in Applied Ethics,” *Ethics*, vol. 116, no. 4, 2006, pp. 656-679.

Horowitz, Tamara, “Philosophical Intuitions and Psychological Theory,” *Ethics*, vol. 108, no. 2, 1998, pp. 367-385.

Kamm, F. M., “Moral Intuitions, Cognitive Psychology, and the Harming-Versus-Not-Aiding Distinction,” *Ethics*, vol. 108, no. 3, 1998, pp. 463-488.

Kamm, F. M., *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*, Oxford University Press, 2007.

Kahneman, Daniel, Knetsch, Jack L. and Thaler, Richard H., “Fairness as a Constraint on Profit-Seeking: Entitlements in the Market,” *American Economic Review*, vol. 76, 1986, pp. 728-741.

Kahneman, Daniel, Knetsch, Jack L. and Thaler, Richard H., “Experimental Tests of the Endowment Effect and the Coase Theorem,” *Journal of Political Economy*, vol. 98, 1990, pp. 1325-1348.

Kahneman, Daniel, Knetsch, Jack L. and Thaler, Richard H., “Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias,” *Journal of Economic Perspectives*, vol. 5, no. 1, 1991, pp. 193-206. Reprinted in Kahneman & Tversky 2000, ch. 8.

[10] 現状維持バイアスの考慮から得られる確かな教訓は、新たな科学技術の導入に際する制度設計の設計段階での検討の重要性である。ひとまず出来合いの制度を導入してみて様子を見ながらより適切なものに微調整していくという、一見して賢明そうな方策が常にうまくいくとは限らない。この後に本文で述べられるように、制度設計の前段階において、準拠点を異にする者たちによる実質的なコラボレーションが行われることは、わが国が得意とするなし崩し型対応では立ち行かない場合に採られるべき有望な方策たりうるだろう。

[11] もちろん、コラボレーションに際して、集団に固有のバイアスの数々を免れないであろう。これに関する考察は機会を改める他はないが、たとえば、必ず社会認識論の研究者を立ち会わせて、同時にそのコラボレーション自体をメタレベルで研究してもらう、といった方策も面白いかもしれない。

Kahneman, Daniel and Tversky, Amos, *Choices, Values, and Frames*, Cambridge University Press, 2000.

Korobkin, Russell, “The Endowment Effect and Legal Analysis,” *Northwestern University Law Review*, vol. 97, no. 3, 2003, pp. 1227-1291.

Samuelson, William and Zeckhauser, Richard, “Status Quo Bias in Decision Making,” *Journal of Risk and Uncertainty*, vol. 1, 1988, pp. 7-59.

Sandel, Michael J., “The Case against Perfection,” *The Atlantic Monthly*, vol. 293, no. 3, April 2004. （Online version http://www.theatlantic.com/doc/200404/sandel）

Sandel, Michael J., *The Case against Perfection: Ethics in the Age of Genetic Engineering*, Belknap Press of Harvard University Press, 2007.

Thaler, R., “Toward a Positive Theory of Consumer Choice,” *Journal of Economic Behavior and Organization*, vol. 1, 1980, pp. 39-60. Reprinted in Kahneman & Tversky 2000, ch. 15.

Tversky, Amos and Kahneman, Daniel, “The Framing of Decisions and the Psychology of Choice,” *Science* vol. 211, no. 4481, 1981, pp. 453-58.

van Roojen, Mark, “Reflective Moral Equilibrium and Psychological Theory,” *Ethics*, vol. 109, no. 4, 1999, pp. 846-857.